DOCUMENT RESUME

ED 437 853 FL 026 129

AUTHOR Joritz-Nakagawa, Jane
TITLE A Cooperative Performance Test for Japanese University Conversation Classes.
PUB DATE 1999-00-00
NOTE 11p.; In: Cooperative Learning. JALT Applied Materials; see FL 026 115.
PUB TYPE Reports - Descriptive (141)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Classroom Techniques; *Communicative Competence (Languages); *Cooperative Learning; *English (Second Language); Foreign Countries; Higher Education; *Language Tests; *Oral Language; Second Language Instruction; Second Language Learning; Teaching Methods; Testing
IDENTIFIERS Japan

ABSTRACT

This document discusses a test devised for a university level English-as-a-foreign-language (EFL) course that tests students' conversational ability. The test has been used successfully in a Japanese university English communication class over three years. Learning not only to speak but to listen and respond appropriately to others is also an important measure of how well one communicates orally in a foreign language. In brief, the test requires students in groups to prepare and practice a conversation evaluated by the teacher. It is a performance assessment, by definition a type of alternative assessment, requiring students to demonstrate knowledge and skill. This test might be difficult to manage in a class with more than 50 students. Students are divided into groups of four and choose a topic (to be approved by the teacher). They then have a few weeks to prepare a 20 to 30 minute discussion. Students were not allowed any use of notes, papers, books, dictionaries, or aids other than visual aids such as photographs, pictures, maps, or props such as jump ropes, CDs, and musical instruments to be used to facilitate discussion on the test day. The test has been well-received by students and meets all six criteria for effective assessment described by Good and Brophy (1994). (An appendix with a test comment form in included.) (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Gene van Troyer
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

Chapter 14

ED 437 853

A Cooperative Performance Test for Japanese University Conversation Classes

Jane Joritz-Nakagawa

Miyagi University

When I first began teaching English conversation courses in Japanese universities, I wasn't sure what kind of test to offer the class. My courses mostly emphasized pair and/or group speaking activities. A fair test, I felt, should reflect the course content and method of instruction.

When I asked colleagues about the kinds of tests they gave, most said they gave paper-and-pencil tests of some kind, often with the reason that class size made it impractical to test student speaking abilities. Some gave tests which involved reading or writing or objective tests which tested only listening ability. However, these tests seemed unrelated to what I was asking students to do in class every week, and therefore unfair. Other teachers concluded the term by having students individually get in front of the class to deliver a speech. While at least speaking was being tested, my courses focussed on having students learn how to appropriately converse with each other in English in class. A speech didn't test such abilities as turn-taking, offering appropriate responses as a listener, etc.

Joritz-Nakagawa, J. (1999). A cooperative performance test for Japanese university conversation classes. In D. Kluge, S. McGuire, D. Johnson, & R. Johnson (Eds.), *JALT applied materials: Cooperative learning* (pp. 179-188). Tokyo: Japan Association for Language Teaching.

FL C26 129

I wanted a test that tested students' conversational ability. I wanted them to not only speak but to listen and respond appropriately to others, as responding appropriately is also an important measure of being able to communicate successfully orally in a foreign language. With the above goals in mind, I eventually came up with the test I will describe in this paper. Realizing that the success of a conversation or discussion hinges on the cooperation of its participants and that the group of participants as a whole is responsible for its success, I quickly decided on a cooperative learning format (group project format) for this test.

In brief, the test requires students in groups to prepare and practice a conversation evaluated by the teacher. It is a performance assessment, by definition a type of alternative assessment, requiring students to demonstrate knowledge and skill (a definition of performance assessment may be found in Choate [1995] or Good & Brophy [1995]).

Overview of the Courses in Which the Test Has Been Used

I have used this test in all of the courses I taught over a three year period at a four year private university in which oral English communication was the focus of the course. In total, there were five different courses, offered at the first, second, and third academic years in which I used this test. Class size ranged from approximately 20 to 50 students. Courses met once a week for 90 minutes throughout the academic year.

This test is appropriate, in my opinion, for any similar *eikaiwa* [conversation] course where the focus has been student pair and/or group speaking activities. However, a teacher teaching with a class size of over 50 will have to consider how to remedy scheduling difficulties (see below) and will need to set the appropriate stage for learning before undertaking the test.

Overview of the Test

Usually, about six or seven weeks before the end of the semester, preparation for the test began. Three weeks were needed to do all of the following: explain the test to the class, get them into groups (these first two generally take half a period or less) and give them in-class time (about 2 1/2 periods) to work on preparing their test. An additional two to three weeks was usually needed for testing.

After explaining the basic requirements for the test, including the criteria for working together and the grading criteria (explained further below) each group chose a topic (one previously discussed in class or a new topic) to discuss for the test. The topics were approved by me. On the test day, each group was instructed to have a 20 to 30 minute discussion. In general, I required a four

member group of first year students to converse for 20 minutes; a second year student group of the same size 20 to 25 minutes, and third year students 25 to 30 minutes.

My class size ranged from 20 to 48, and students were generally grouped into four member groups. Therefore, the number of groups per class tended to be between 5 and 12. Teachers with larger enrollments may need to make scheduling adjustments. If instructional days are not to be reduced further, non-instructional time (at the beginning or ending of the day, lunch hours, etc.) would have to be used for testing. Videotaping or audiotaping the tests outside of class may be possible, but requires minimally that students do this outside of class and also perhaps additional teacher time to view or listen to the tapes outside of class, as well as access to equipment. Other options could be reducing the test time or increasing group size. In cases where the teacher needs to shorten the test time, the difficulty level of the topic could perhaps be increased. Teachers who schedule the test during instructional periods but are concerned about missed instructional days could ask students to do other work outside of class to make up the time, assuming students do not need to come to class for all the testing days. Alternately students could be required to attend (though I did not choose this option for reasons I will explain below).

Students were not allowed to use any notes, papers, books, dictionaries, or aids other than visual aids such as photographs, pictures, or maps used to facilitate discussion on the test day. Some students whose topic was, for example, foreign travel, used travel photographs and area maps which they showed their group members as they conversed. Other aids used artfully by students included high school year books, family photographs, charts or pictures they drew on the black board, Japanese magazines which they talked about in English, and miscellaneous paraphernalia and props including such things as jump ropes, CDs, and guitars. Students were very creative in their use of aids; these made the conversations both more lively and more comprehensible. Texts were disallowed so that there was no reading from a paper.

The full grading criteria (explained below) were given to students at the time the test was explained, immediately preceding grouping the class for the test.

Grouping the Students into Test Groups

I have used three methods for grouping the class:

a. The students were free to sign up for any test day and time available;
b. The grouping was the result of chance, where students choosing a card with a letter or number on it determined which group they were to work with;
c. I have on one occasion hand picked the groups using as my criteria attendance patterns (those with a similar rate of attendance were grouped to-

to rise to the highest level of combined strengths in the group. Since I prefer the test to be a learning tool (doing double duty as an evaluation tool), not a separating or screening device, this grouping method appears wanting. Such grouping may bias the teacher just at the time she is engaged in the task of assigning grades to students. Further, a positive class experience while working on this test may encourage so-so contributors to contribute more in the future to the course. I believe this grouping method was one of my early mistakes in learning myself to use cooperative learning. (Another was assigning individual grades to each participant but it was impossible to extricate the individual contributions from the group performance, and I feel this method does not reflect the teacher's commitment to the idea that the group must be accountable as a group.)

Presenting the Test Objectives and Criteria to Students

I gave students a handout which explained the basic information they needed to know to prepare for the test. Students read this handout. Then, it was explained by me orally. Finally, there was a question and answer period about the objectives and guidelines. There were few if any questions and answers because students already knew about the test, as it was explained at the onset of the course, and was not a radical departure from our regular instruction. (See the Appendix A for a sample of a handout given to a second year oral English class.)

Group Preparation for the Test

As mentioned above, each group met usually for 2 $^{1/2}$ class periods to discuss what topic they will discuss, decide the content, and practice their discussion for the test. In my experience students request and seem to want very little help from me during this time. The sort of help I gave, upon request, was usually restricted to answering language questions (often, checking the grammaticality of questions or confirming word usage as appropriate). Most groups wished to use all of the allotted time to prepare; occasionally I had a group who said they had finished preparing early (e.g,, in 1 $^{1/2}$ or 2 class periods). I brought along work for students to do (usually something such as a conversation board game or vocabulary game) just in case some finished their work early. In this instance, I allowed those groups who wished to prepare to do so, and others could take from me an activity (or devise their own activity if appropriate).

Most students approached preparation for the test with visible diligence and seriousness; in fact, the vast majority of students appeared to wholeheartedly enjoy preparing for the test. The classroom atmosphere tended to be extremely animated during this time, often punctuated by a lot of laughter. (Many students incorporated humor into their discussion.) Occasionally I had some groups who seemed to be wasting their class preparation time though these were the exception to the rule. Upon observing this, I approached the group and at-

gether) and secondarily where a choice remained (e.g., 12 students all have perfect attendance), and diversity (e.g., creating gender diverse groups).

Results of the Three Methods

Method A. Being free to sign up for any test day and time available, which allows students to choose the membership of their group, has been the overwhelming favorite of second and third year student classes, particularly smaller classes. When asked to pick a method from among the three described above, a show of hands of all classes but one showed 80 to 90+% of second and third year students in favor of this method.

Method B. The grouping by chance (lottery) method has been the overwhelming favorite (usually about 90% in favor) of the first year students, who also were asked to vote for the grouping method they preferred through a show of hands. I believe first year students prefer this method because, unlike second and third year students, they don't know the others in the class as well socially. With cliques being less defined in first year perhaps it is uncomfortable for them to choose whom to work with.

Method C. In the one second year class which voted to have me to make the groups rather than them, I decided (and they concurred) that going by attendance record in the class was a fair method of grouping them. Percentage of classes attended was the first criterion and where that left further options the secondary criterion was diversity (gender, nationality, or age) with the idea, as Maznevski (1994) has noted, that diversity improves group performance. However, I think I will probably not repeat this method. The reason I disliked this method was that it seemed that performance on the test did turn out to be directly correlated with the attendance patterns; in other words, those with high or perfect attendance ended up with A's on the test, and those with so-so attendance got so-so marks, etc. I wondered if, as is often the case with ability groupings, the results were based on a kind of self-fulfilling prophesy. I thought at first this method was fair, because it seemed unfair to put a diligent student in a group, for example, who had never missed a lesson with a "lazy" one who had perhaps missed nearly half of the lessons. Certainly, the latter can pose a problem to be worked out, but working out such problems is part of the work of a cooperative learning class. Trying to shield students from such problems may diminish their growth experience. (Of course laziness is not the only explanation for student absence.)

Diversity of any kind can be beneficial to all in terms of offering a balance of strengths and weaknesses and allowing members to positively influence each other. (See Joritz-Nakagawa, 1997, for a brief look at the advantages of heterogeneity and dangers of self-fulfilling prophecies.) In my experience, and as has been noted above to be characteristic of high-functioning groups, most often the group seems

tempted to see what was going on by asking questions, and tried to encourage them toward working diligently on test preparation if they didn't seem to be working hard. In these cases I might help resolve a problem if that were the reason for the interruption in work or ask the group detailed questions about their topic to identify areas where they could further prepare.

On the test day it should be clear if any groups haven't worked diligently as the quality of their work will generally be substantially lower than other groups. The exception may be if you have a very broad range of levels in your class including students with very high facility in English who perhaps could do well on this type of test without expending a lot of effort. However, I feel while it is possible for anyone to succeed in this test, it is also unlikely for even an especially talented individual or group to function well without preparing somewhat seriously, and especially preparing together, for this test.

In short, individual effort and group cooperation are the key determinants of success. Discussions of groups who haven't prepared thoroughly will have a random character to them; the pace of discussion will be adversely affected; there will be many more hesitations, for example when groping for something to say. If they didn't practice well together, what members say may not flow together well (will sound disjointed, etc.). These qualities lower the grade on the test and may prevent passing, as the test and grading criteria explained above shows. Students are instructed to speak naturally but the discussion is supposed to come off more like a panel discussion on TV, where the participants know in advance what the topic is and have taken some pains to prepare; which enables them to participate smoothly, intelligently, and effectively. However, students are advised not to sound overly rehearsed. (See test instructions and criteria explained above.)

The Examinations

At the time of this writing, I have given this test on over 30 occasions. The first time I administered this test to student groups in five different courses, I was overwhelmed at seeing what the students could do. I had been particularly worried that many of my first year students, for example, couldn't carry on this type of 20 minute discussion successfully, but this was not the case. The success rate continued with the only disappointments, and these have been few, being students who had ignored the instructions for the test and didn't prepare sufficiently. I gave this test both at midterm and at the end of the academic year, and have found the few groups who test poorly at midterm tend to make amends at the final. However it may be possible that students who do very well at midterm could slack off at the final for full year courses if they are not highly motivated or depending upon the weight of the test in terms of the final grade, for example. Though I gave a weighting of half of the course grade to this test, the teacher needs to determine what weighting in her situation makes sense.

Another pleasure has been to witness the teamwork evident in mixed level groupings. If one student gets stuck during the test, it is usual that another member will assist by, for example, supplying the word or sentence that the other can't recall or hasn't in fact learned, or will clarify what the other has said if they have communicated in such a way that the meaning might be unclear (e.g., A: *Do you mean . . . .?* B: *Yes, yes!*).

I have also seen the varying of roles to complement the varying strengths of members, for example, through the group's choice of a discussion leader or "emcee." Of course these behaviors are in the group's best interest as they are being evaluated as a group based on such criteria as everyone appearing competent, unimpeded flow of communication and so forth, but the teacher will also see it occur naturally in groups with harmonious relations whether or not any formal evaluation is being done. In other words, I view this as the natural result of groups with, to use Goleman's (1995) term, a high group IQ.

The test need not force the cooperation but rather test criteria should be consistent with the practices that have been valued in the classroom. There is no greater experience I think to seeing your students, one by one, group by group, engaging in effective and frequently stimulating conversations done entirely in English without reaching for a set of notes or a dictionary, and helping each other be effective. Interestingly, I have witnessed numerous students practicing with their groups on their own time outside of class. Since many students otherwise appear loathe to do homework, this seems to me a good sign indeed.

Occasionally, some students will balk at the test when I first explain it, thinking "I can't do this!" Succeeding gives students a confidence in their abilities—this confidence may be the best gift you can give them. Success on the midterm test appears to often lead to increased confidence in speaking during the second semester.

One other difficulty for the teacher may be, if she has many classes in which this test will be given, the two or three test weeks will actually take a lot of her energy. While there is little preparation for the teacher (in fact the last five or six weeks of the term, following the schedule I have outlined here based on my class size and school calendar, required little preparation on my part), listening attentively to group after group and being, of course, responsible for evaluating the students while doing so requires some resilience. But the popularity of this test and my feeling that it is appropriate and fair given the teaching situation makes it, I think, worthwhile.

Grading and Feedback

As mentioned above, in my oral English courses student course grades for each term were calculated as 50% group test grade and 50% individual participation grade. The group test grade reflected my impression of how well the group

fulfilled the criteria described above. The participation grade reflected the percentage of classes students individually attended, and also took into account factors such as tardiness and level of participation (class behavior). A student who attended all courses in the term and participated adequately each time would thus receive 100 points for their participation grade.

Students received their individual participation grades during a private conference with the teacher, and the group grade during a group conference with the teacher. Students were encouraged to give their own impression of their work during these conferences. Students also did anonymous evaluations of the course as a whole usually during the first semester, and often again in the second semester. These evaluations solicited such information as which topics they were interested in discussing, their feeling about the balance between listening and speaking activities, and what size groups they preferred for conversation.

I gave each group a copy of a test feedback form (see Appendix). I merely circled strong/weak points; this made it easy for me to complete the form and easy for the group to understand it. The form can be explained to the whole class before giving the feedback. This information is not new but is the same as the test instructions and criteria students received before the test, explained above, but in a different form. Any additional comments were written in the margins or on the bottom or reverse side of the form. Additional comments were, for example, responses to specific ideas from or comments about specific language used in the discussion.

At the end of the list I have mentioned vocabulary, structure, pronunciation, and intonation. However, these elements were considered important for purposes of grading the group only as far as they (e.g., word choice/grammar/pronunciation) might inhibit comprehensibility of the discussion or diminish its impact (e.g., consistently flat intonation, though perhaps indicative of lack of English mastery, may come off as lack of enthusiasm).

Summary

In general, I have been abundantly pleased observing the process of student groups working together on this test as well as their finished products. Students frequently commented that they enjoyed this test. Some having taken the test at midterm have asked me: Could we please have this same test again at the final? While test grades have varied, only those who have ignored the test instructions have failed the test; it is designed so that any group of individuals, with effort and cooperation skills, have the possibility of passing the course.

This test also meets important criteria for performance tests as described by Good and Brophy (1994, p. 641): 1) students should be assigned assessment tasks that are educative and engaging (i.e., not just memorize lists); 2) students should

know what the performance standards are (they should be subjected to minimal secrecy in testing and grading); 3) there should be clear criteria for grading work; 4) students need ample opportunity (e.g., enough time) to produce successful work; 5) students should be given the opportunity to display and document their positive achievements (versus tests which aim to reveal their weaknesses); and 6) students are given the chance to question grades and test practices.

I feel this test is an appropriate and fair conclusion to a semester of pair and group conversation in the Japanese university setting, where care has been taken to create an atmosphere conducive to cooperative learning and schedule the activities appropriately. Students participating in the test are engaged in a learning experience which requires them to be responsible for creating and carrying out a discussion in a group.

References

Goleman, Daniel. (1995). *Emotional intelligence*. New York: Bantam.

Good, T. L. & Brophy, J. 1995. *Contemporary educational psychology* (5th ed.). New York: Longman.

Johnson, D.W., Johnson, R., & Smith, K. (1991). *Active learning: Cooperation in the college classroom*. Edina: MN: Interaction Book Company.

Joritz-Nakagawa, J. (1997). Track and stream. *College and University Educators' Newsletter 5*(1).

Kagan, S. (1994). *Cooperative learning*. San Clemente, CA: Resources for Teachers, Inc.

Maznevski, M. (1994). Understanding our differences. *Human Relations, 47*(5), 531-552.

Appendix

Test Comment Form

Test Comment Form
Class: Listening Speaking Discussion I II
Day: Tu We Fr

Instructor: J. Nakagawa
English Convers. I
Per: 1 2 3

Group number:_____
Member names:_____
Test grade:_____ (= 50% of TERM grade)

Impression of strong points:

_____ Easy to understand
_____ Loud enough
_____ Interesting /stimulating topic

_____ Talked about only one topic
_____ Appropriate length
_____ Shared time equally
_____ Natural sounding
_____ Sounded relaxed / good pace
_____ Used English "aizuchi" and Q&A
_____ Seemed to enjoy conversing
_____ Adequately prepared
_____ Spoke without notes
_____ On time for test and generally followed instructions
_____ Good use of English vocabulary and structure
_____ Good pronunciation / intonation
_____ Used only English

Impression of weak points:

_____ Sometimes/often hard to understand
_____ Sometimes/often couldn't hear
_____ Content inappropriate/too easy/go into more depth
_____ Talked about more than 1 topic
_____ Too short
_____ Did not share time equally
_____ Sounded rehearsed/unnatural
_____ Seemed nervous/ slow paced
_____ Need to use "aizuchi"/Q&A more
_____ Didn't show a lot of enthusiasm
_____ Preparation seemed inadequate
_____ Tried to use notes
_____ Were late for test and/or didn't follow instructions
_____ English vocab./structure below average for this academic year
_____ Need to work on pron./intonation
_____ Occasional use of native languages